The Bivariate Normal Distribution
The Two Distributions Are Correlated

Gary Schurman, MBE, CFA

July, 2023

The bivariate normal distribution is used to determine the **joint probability** of pulling a random variate with a given value from one normal distribution and pulling another random variate with a given value from another normal distribution. To have a bivariate normal distribution or a joint normal distribution, both random variables need to be normally distributed and independent.

In this white paper we will develop the mathematics for the bivariate normal distribution and then use those equations to calculate joint probabilities for a hypothetical case where the distribution of random variable x and random variable y are correlated.

Our Hypothetical Problem

The table below summarizes the parameters to our problem...

μ_x	=	The mean of the random variate x	=	6.00
σ_x	=	The standard deviation of the random variate x	=	5.00
μ_y	=	The mean of the random variate y	=	4.00
σ_y	=	The standard deviation of the random variate y	=	3.00
ρ	=	Correlation coefficient	=	0.80

The problem: The random variate a is pulled from the distribution of x and the random variate b is pulled from the distribution of y. Using the table above, what is the joint probability that the random variate pulled from the distribution of x is less than 4.5 **and** the random variate pulled from the distribution of y is less than 3.5?

Ordinary Least Squares Estimation

We will define the variable ρ to be the correlation of the two normal probability distributions of x and y. The definition of correlation is...

$$\rho = \text{Correlation of random variable } x \text{ and random variable } y \tag{1}$$

To introduce correlation into the bivariate normal distribution we will model the normally-distributed random variable y (dependent variable) to be a function of the normally-distributed random variable x (independent variable). Our ordinary least squares estimate (i.e. linear regression) of the value of y given the value of x is... [2]

$$y = \alpha + \beta\, x + \epsilon \tag{2}$$

The equations for the linear regression variables α and β in Equation (2) above are... [2]

$$\alpha = \mu_y - \beta\, \mu_x \;\; \text{...and...} \;\; \beta = \rho\, \frac{\sigma_y}{\sigma_x} \tag{3}$$

The equations for the expected values of the error term in Equation (2) above are... [2]

$$\mathbb{E}\Bigg[\epsilon\Bigg] = 0 \;\; \text{...and...} \;\; \mathbb{E}\Bigg[\epsilon^2\Bigg] = (1 - \rho^2)\, \sigma_y^2 \tag{4}$$

Building Our Model

When correlation is zero the equations for the probability density functions $f(x)$ and $g(y)$ from Part I are... [1]

$$f(x) = \sqrt{\frac{1}{2\,\pi\,\sigma_x^2}}\,\mathrm{Exp}\left\{-\,\frac{1}{2}\,\frac{(x-\mu_x)^2}{\sigma_x^2}\right\} \text{ ...and... } g(y) = \sqrt{\frac{1}{2\,\pi\,\sigma_y^2}}\,\mathrm{Exp}\left\{-\,\frac{1}{2}\,\frac{(y-\mu_y)^2}{\sigma_y^2}\right\} \tag{5}$$

To introduce correlation we will redefine the second probability density function ($g(y)$) in Equation (5) above as...

$$\hat{g}(y) = \sqrt{\frac{1}{2\,\pi\,\hat{\sigma}_y^2}}\,\mathrm{Exp}\left\{-\,\frac{1}{2}\,\frac{(y-\hat{\mu}_y)^2}{\hat{\sigma}_y^2}\right\} \text{ ...where... } \hat{\mu}_y \neq \mu_y \text{ ...and... } \hat{\sigma}_y^2 \neq \sigma_y^2 \tag{6}$$

In the no correlation scenario, the value of the random variable y is drawn from the independent, normal distribution of y. We will define the variable $\hat{y}$ to be the replacement for y in the correlation scenario. Using Appendix Equation (13) below, the equation for $\hat{y}$ is...

$$\hat{y} = \mu_y + \rho\,\frac{\sigma_y}{\sigma_x}\,(x-\mu_x) + \epsilon \tag{7}$$

Using Appendix Equation (14) below, the equation for the mean of $\hat{y}$ in Equation (7) above is...

$$\hat{\mu}_y = \mathbb{E}\Big[\hat{y}\Big] = \mu_y + \rho\,\frac{\sigma_y}{\sigma_x}\,(x-\mu_x) \tag{8}$$

Using Appendix Equations (15) and (16) below, the equation for the variance of $\hat{y}$ in Equation (7) above is...

$$\hat{\sigma}_y^2 = \mathbb{E}\Big[\hat{y}^2\Big] - \Big[\mathbb{E}\Big[\hat{y}\Big]\Big]^2 = (1-\rho^2)\,\sigma_y^2 \tag{9}$$

Using Equations (8) and (9) above, we can rewrite Equation (6) above as...

$$\hat{g}(y) = \sqrt{\frac{1}{2\,\pi\,(1-\rho^2)\,\sigma_y^2}}\,\mathrm{Exp}\left\{-\,\frac{1}{2}\,\frac{(y-\mu_y-\rho\,\frac{\sigma_y}{\sigma_x}\,(x-\mu_x))^2}{(1-\rho^2)\,\sigma_y^2}\right\} \tag{10}$$

Using Equations (5) and (10) above, the equation for the joint probability that x is less than a and y is less than b when the distributions of random variable x and random variable y are correlated is... [1]

$$\begin{aligned}\mathrm{Prob}\Big[x \leq a, y \leq b\Big] &= \int\limits_{-\infty}^{a}\int\limits_{-\infty}^{b}\sqrt{\frac{1}{2\,\pi\,\sigma_x^2}}\,\mathrm{Exp}\left\{-\,\frac{1}{2}\,\frac{(x-\mu_x)^2}{\sigma_x^2}\right\}\sqrt{\frac{1}{2\,\pi\,(1-\rho^2)\,\sigma_y^2}}\,\mathrm{Exp}\left\{-\,\frac{1}{2}\,\frac{(y-\mu_y-\rho\,\frac{\sigma_y}{\sigma_x}\,(x-\mu_x))^2}{(1-\rho^2)\,\sigma_y^2}\right\}\delta y\,\delta x\\ &= \int\limits_{-\infty}^{a}\int\limits_{-\infty}^{b}\frac{1}{2\,\pi}\sqrt{\frac{1}{(1-\rho^2)\,\sigma_x^2\,\sigma_y^2}}\,\mathrm{Exp}\left\{-\,\frac{1}{2}\left[\frac{(x-\mu_x)^2}{\sigma_x^2}+\frac{(y-\mu_y-\rho\,\frac{\sigma_y}{\sigma_x}\,(x-\mu_x))^2}{(1-\rho^2)\,\sigma_y^2}\right]\right\}\delta y\,\delta x\end{aligned} \tag{11}$$

The Answer To Our Hypothetical Problem

We will use numerical integration to approximate the integral in equation (11) above. **The VBA code for numerical integration is included at the end of this white paper**.

The probability of x being less than or equal to 4.5 and y being less than or equal to 3.5 given a correlation coefficient of 0.80 is

$$\mathrm{Prob}\Big[x \leq 4.5, y \leq 3.5\Big] = \mathrm{BVNProb}(4.5,\ 3.5,\ 6,\ 4,\ 25,\ 9,\ 0.80) = 0.3069 \tag{12}$$

Appendix

A. Using Equations (2) and (3) above, the equation for the random variable $\hat{y}$ is...

$$\begin{aligned}\hat{y} &= \alpha + \beta\,x + \epsilon \\ &= \mu_y - \beta\,\mu_x + \beta\,x + \epsilon \\ &= \mu_y - \rho\,\frac{\sigma_y}{\sigma_x}\,\mu_x + \rho\,\frac{\sigma_y}{\sigma_x}\,x + \epsilon \\ &= \mu_y + \rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x) + \epsilon \end{aligned} \tag{13}$$

B. Using Equations (4) and (13) above, the equation for the expected value of the random variable $\hat{y}$ is...

$$\begin{aligned}\mathbb{E}\bigg[\hat{y}\bigg] &= \mathbb{E}\bigg[\mu_y + \rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x) + \epsilon\bigg] \\ &= \mu_y + \rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x) + \mathbb{E}\bigg[\epsilon\bigg] \\ &= \mu_y + \rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x) \end{aligned} \tag{14}$$

C. Using Equation (14) above, the equation for the square of the expected value of the random variable $\hat{y}$ is...

$$\bigg[\mathbb{E}\bigg[\hat{y}\bigg]\bigg]^2 = \mu_y^2 + \bigg(\rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x)\bigg)^2\,\mu_y^2 + 2\,\rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x)\,\mu_y \tag{15}$$

D. Using Equations (4) and (13) above, the equation for the expected value of the square of the random variable $\hat{y}$ is...

$$\begin{aligned}\mathbb{E}\bigg[\hat{y}^2\bigg] &= \mathbb{E}\bigg[(\mu_y + \rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x) + \epsilon)^2\bigg] \\ &= \mathbb{E}\bigg[\mu_y^2 + \bigg(\rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x)\bigg)^2 + \epsilon^2 + 2\,\mu_y\,\rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x) + 2\,\mu_y\,\epsilon + 2\,\rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x)\,\epsilon\bigg] \\ &= \mu_y^2 + \bigg(\rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x)\bigg)^2 + \mathbb{E}\bigg[\epsilon^2\bigg] + 2\,\mu_y\,\rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x) + 2\,\mu_y\,\mathbb{E}\bigg[\epsilon\bigg] + 2\,\rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x)\,\mathbb{E}\bigg[\epsilon\bigg]\bigg] \\ &= \mu_y^2 + \bigg(\rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x)\bigg)^2 + (1 - \rho^2)\,\sigma_y^2 + 2\,\rho\,\frac{\sigma_y}{\sigma_x}\,(x - \mu_x)\,\mu_y \end{aligned} \tag{16}$$

References

[1] Gary Schurman, *The Bivariate Normal Distribution - No Correlation,* July, 2023.

[2] Gary Schurman, *Univariate Ordinary Least Squares Estimator,* May, 2011.

Listing 1: VBA Code

```
'Name: BVNProb
'Purpose: Probability that random variable x < a and random variable y < b.
'Author: Gary Schurman, MBE, CFA
Public Function BVNProb(a As Double, b As Double, mean_x As Double, mean_y As Double, _
variance_x As Double, variance_y As Double, rho As Double) As Double
    'Declare calculation variables.
    Dim mValue As Double
    'Define calculation variable values.
    If rho = 0 Then
        mValue = BVNProbUncorrelated(a, b, mean_x, mean_y, variance_x, variance_y)
    Else
        mValue = BVNProbCorrelated(a, b, mean_x, mean_y, variance_x, variance_y, rho)
    End If
    'Return value to caller.
    BVNProb = mValue
End Function

'Name: BVNProbUncorrelated
'Purpose: Returns probability that x < a and random variable y < b. Rho is zero.
'Author: Gary Schurman, MBE, CFA
Private Function BVNProbUncorrelated(a As Double, b As Double, mean_x As Double, _
mean_y As Double, variance_x As Double, variance_y As Double) As Double
    'Declare calculation variables.
    Dim mValue As Double
    Dim mValue01 As Double
    Dim mValue02 As Double
    'Define calculation variable values.
    mValue01 = CNDF(a, mean_x, variance_x)
    mValue02 = CNDF(b, mean_y, variance_y)
    mValue = mValue01 * mValue02
    'Return value to caller.
    BVNProbUncorrelated = mValue
End Function

'Name: BVNProbCorrelated
'Purpose: Returns probability that x < a and random variable y < b. Rho is non-zero.
'Author: Gary Schurman, MBE, CFA
Private Function BVNProbCorrelated(a As Double, b As Double, mean_x As Double, _
mean_y As Double, variance_x As Double, variance_y As Double, rho As Double) As Double
    'Declare calculation variables.
    Dim mValue As Double
    Dim mALowerBound As Double
    Dim mAUpperBound As Double
    Dim mBLowerBound As Double
    Dim mBUpperBound As Double
    Dim mStepSize As Double
    Dim mNumberSteps As Integer
    Dim mPDF As Double
    Dim x As Double
    Dim y As Double
    Dim dx As Double
    Dim dy As Double
    'Define integral parameters.
    mValue = 0
    mNumberSteps = 1000
```

```
    mAUpperBound = a
    mBUpperBound = b
    mALowerBound = mean_x - Sqr(variance_x) * 4
    mBLowerBound = mean_y - Sqr(variance_y) * 4
    dx = (mAUpperBound - mALowerBound) / (mNumberSteps - 1)
    dy = (mBUpperBound - mBLowerBound) / (mNumberSteps - 1)
    'Calculate joint probability.
    For x = mAUpperBound To mALowerBound Step -dx
        For y = mBUpperBound To mBLowerBound Step -dy
            mPDF = CBVNormalPDF(x, y, mean_x, mean_y, variance_x, variance_y, rho)
            mValue = mValue + mPDF * dx * dy
        Next y
    Next x
    'Return value to caller.
    BVNProbCorrelated = mValue
End Function

'Name: CBVNormalPDF
'Purpose: Returns the bivariate-normal distribution probability density.
'Author: Gary Schurman, MBE, CFA
Private Function CBVNormalPDF(a As Double, b As Double, mean_a As Double, mean_b As Double, _
variance_a As Double, variance_b As Double, rho As Double) As Double
    'Declare calculation variables.
    Dim mValue As Double
    Dim mValue01 As Double
    Dim mValue02 As Double
    Dim mValue03 As Double
    Dim mValue04 As Double
    Dim mValue05 As Double
    Dim mValue06 As Double
    'Define calculation variable values.
    mValue01 = 1 / (2 * PIValue)
    mValue02 = (1 - rho ^ 2) * variance_a * variance_b
    mValue03 = (a - mean_a) ^ 2
    mValue04 = (b - mean_b - rho * Sqr(variance_b / variance_a) * (a - mean_a)) ^ 2
    mValue05 = variance_a
    mValue06 = (1 - rho ^ 2) * variance_b
    mValue = mValue01 * Sqr(1 / mValue02)
    mValue = mValue * Exp(-1 / 2 * (mValue03 / mValue05 + mValue04 / mValue06))
    'Return value to caller.
    CBVNormalPDF = mValue
End Function

'Name: CNDF
'Purpose: Return cumulative normal distribution function.
Private Function CNDF(z As Double, mean As Double, variance As Double) As Double
    CNDF = Application.WorksheetFunction.NormDist(z, mean, Sqr(variance), True)
End Function

'Name: PIValue
'Purpose: Returns the value of pi.
Private Function PIValue() As Double
    PIValue = Application.WorksheetFunction.Pi()
End Function
```

.